

University of Montana

## ScholarWorks at University of Montana

---

Graduate Student Theses, Dissertations, &  
Professional Papers

Graduate School

---

1953

### Discussion of quadratic forms in normally distributed random variables

Thomas Arthur Bray  
*The University of Montana*

Follow this and additional works at: <https://scholarworks.umt.edu/etd>

## Let us know how access to this document benefits you.

---

#### Recommended Citation

Bray, Thomas Arthur, "Discussion of quadratic forms in normally distributed random variables" (1953).  
*Graduate Student Theses, Dissertations, & Professional Papers*. 8235.  
<https://scholarworks.umt.edu/etd/8235>

This Thesis is brought to you for free and open access by the Graduate School at ScholarWorks at University of Montana. It has been accepted for inclusion in Graduate Student Theses, Dissertations, & Professional Papers by an authorized administrator of ScholarWorks at University of Montana. For more information, please contact [scholarworks@mso.umt.edu](mailto:scholarworks@mso.umt.edu).

A DISCUSSION OF QUADRATIC FORMS IN NORMALLY  
DISTRIBUTED RANDOM VARIABLES

by

THOMAS A. BRAY


B. A. Montana State University, 1952

Presented in partial fulfillment of the requirements  
for the degree of  
Master of Arts

Montana State University

1953

Approved by:

  
Chairman, Board of Examiners

  
Dean, Graduate School

  
Date

UMI Number: EP39036

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI EP39036

Published by ProQuest LLC (2013). Copyright in the Dissertation held by the Author.

Microform Edition © ProQuest LLC.

All rights reserved. This work is protected against  
unauthorized copying under Title 17, United States Code



ProQuest LLC.  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

# ACKNOWLEDGEMENT

I am indebted to Professor George Marsaglia who has generously permitted me to use some of the results of his research.

T. A. B.

## TABLE OF CONTENTS

INTRODUCTION . . . . .	1
CHAPTER	
I. SELECTED THEOREMS FROM THE THEORY OF MATRICES . . . .	2
II. SELECTED THEOREMS FROM THE THEORY OF MATHEMATICAL STATISTICS . . . . .	11
III. QUADRATIC FORMS IN NORMALLY DISTRIBUTED RANDOM VARIABLES . . . . .	19
BIBLIOGRAPHY . . . . .	28

## INTRODUCTION

The purpose of this discussion is to present some of the properties of quadratic forms involving normally distributed random variables and some of the theorems from the theory of matrices and the theory of mathematical statistics on which these results are based. A knowledge of matrix theory and mathematical statistics is assumed. The principal results from the theory of matrices appear in Chapter I. Chapter II presents some of the results of mathematical statistics, and Chapter III contains the material on quadratic forms. The theorems are numbered consecutively in each chapter, e. g., the second theorem in the first chapter is designated Theorem 1.2.

With regard to notation, the following system is used:

1. Upper case letters designate square matrices with real elements.
2. Lower case letters designate real numbers.
3. Lower case Greek letters designate real vectors.

Deviations from this notation are specifically noted as they occur.

## CHAPTER I

### SELECTED THEOREMS FROM THE THEORY OF MATRICES

This chapter is devoted to certain theorems from matrix theory which prove useful in the development of Chapter III. The principal result is the well-known theorem on the transformation of a symmetric matrix to diagonal form (Theorem 1.4). This theorem plays such a fundamental role in the theory of quadratic forms that it is presented in detail, even though the result is well-known and may be readily found in most expositions on matrix theory.

A knowledge of the definitions of the null space of a matrix and the minimum function of a matrix is assumed, however, for the sake of clarity these concepts are defined here.

DEFINITION: Let  $A$  be a real matrix and  $m(x)$  be a polynomial with real coefficients. If  $m(x)$  is the polynomial of lowest degree such that  $m(A) = 0$ , then  $m(x)$  is said to be the minimum function of  $A$ .

DEFINITION: The null space of a matrix  $A$  is the set of all vectors  $\xi$  such that  $A\xi = 0$ .

The following three theorems will serve to establish the principal result of this chapter.

THEOREM 1.1: Let  $\beta_1, \dots, \beta_k$  be any  $k$  vectors which span the vector space  $S_k$  of dimension  $k$ . For any value of  $i$ ,  $1 \leq i \leq k$ ,

it is possible to find a vector  $\alpha_1$  such that

$$\beta_1, \dots, \beta_{i-1}, \alpha_1, \beta_{i+1}, \dots, \beta_k$$

span  $S_k$  and such that

$$\beta_j \alpha_1 = 0, \quad j = 1, \dots, i-1.$$

PROOF: Set

$$\alpha_1 = \sum_{h=1}^i l_h \beta_h$$

where the coefficients  $l_1, \dots, l_i$  are to be determined so that

$$\beta_j \alpha_1 = \sum_{h=1}^i l_h \beta_j \beta_h = 0, \quad j = 1, \dots, i-1.$$

The product  $\beta_j \beta_h$  is a real number so the above condition is a system of  $i-1$  linear homogeneous equations in the  $i$  unknowns  $l_1, \dots, l_i$  which has a solution not composed entirely of 0's.

In order to show that the linear systems of vectors

$$\beta_1, \dots, \beta_{i-1}, \beta_k \text{ and } \beta_1, \dots, \alpha_1, \dots, \beta_k$$

are linearly equivalent, it is sufficient to show that  $\beta_i$  can be written as a linear combination of the vectors  $\beta_1, \dots, \beta_{i-1}, \alpha_1$ . That is, it is sufficient to show that  $l_i \neq 0$ .

Suppose  $l_i = 0$ , then

$$\alpha_1 = l_1 \beta_1 + \dots + l_{i-1} \beta_{i-1}.$$

Since  $\alpha_1$  is orthogonal to each of these vectors, it is orthogonal to every linear combination of them and hence to itself. It must, therefore, be the zero vector and hence

$$l_1 \beta_1 + \dots + l_{i-1} \beta_{i-1} = 0.$$

But the vectors  $\beta_1, \dots, \beta_k$  are linearly independent by assumption, so that

$$l_1 = \dots = l_{i-1} = 0$$

and there is a contradiction. Thus,  $l_i \neq 0$  and the theorem is proved.



**THEOREM 1.2:** If  $S_k$  of dimension  $k$  is a subspace of a total vector space  $S$  of dimension  $n$ , then an orthonormal basis

$$\beta_1, \dots, \beta_k, \dots, \beta_n$$

can be found for  $S$  such that

$$\beta_1, \dots, \beta_k$$

is an orthonormal basis for  $S_k$ .

**PROOF:** Let

$$\gamma_1, \dots, \gamma_k, \dots, \gamma_n$$

be a basis for  $S$  such that

$$\gamma_1, \dots, \gamma_k$$

is a basis for  $S_k$ . Let  $\gamma_1 = \alpha_1$ . Then, by Theorem 1.1,  $\gamma_2$  may be replaced by  $\alpha_2$  such that

$$\alpha_1 \alpha_2^i = 0.$$

Similarly,  $\gamma_3$  may be replaced by  $\alpha_3$  such that

$$\alpha_1 \alpha_3^i = \alpha_2 \alpha_3^i = 0.$$

Continuing in this manner,

$$\alpha_1, \dots, \alpha_k$$

forms an orthogonal basis for  $S_k$ , and

$$\alpha_1, \dots, \alpha_k, \dots, \alpha_n$$

forms an orthogonal basis for  $S$ .

Let

$$\alpha_1 = (a_{11}, \dots, a_{1n}) \text{ and } c_1 = \frac{1}{\sqrt{a_{11}^2 + \dots + a_{1n}^2}}.$$

Then

$$\beta_1 = \alpha_1 c_1, \dots, \beta_k = \alpha_k c_k$$

forms an orthonormal basis for  $S_k$  and

$$\beta_1 = \alpha_1 c_1, \dots, \beta_k = \alpha_k c_k, \dots, \beta_n = \alpha_n c_n$$

forms an orthonormal basis for  $S$ .

**THEOREM 1.3:** Let  $A$  be a real symmetric matrix such that the minimum function of  $A$  possesses two distinct linear factors  $l_1(x)$  and  $l_2(x)$ . Let  $\phi_1$  be a vector of the null space of  $l_1(A)$  and  $\phi_2$  be a vector of the null space of  $l_2(A)$ . Then  $\phi_1$  and  $\phi_2$  are orthogonal.

**PROOF:** Let

$$l_1(x) = (x_1 - x) \text{ and } l_2(x) = (x_2 - x)$$

where  $x_1 \neq x_2$ . By definition of the null space of  $l_1(A)$ ,

$$l_1(A)\phi_1^i = 0 \text{ and } A\phi_1^i = x_1\phi_1^i.$$

Similarly,

$$l_2(A)\phi_2^i = 0 \text{ and } A\phi_2^i = x_2\phi_2^i.$$

Now,

$$\phi_2 A \phi_1^i = \phi_2 x_1 \phi_1^i = x_1 \phi_2 \phi_1^i.$$

Since  $A$  is symmetric,  $A = A'$  and

$$\phi_2 A \phi_1^i = (\phi_2 A') \phi_1^i = x_2 \phi_2 \phi_1^i.$$

Consequently,

$$x_1 \phi_2 \phi_1^i = x_2 \phi_2 \phi_1^i$$

and

$$x_1 \phi_2 \phi_1^i - x_2 \phi_2 \phi_1^i = 0.$$

Factoring,

$$(x_1 - x_2) \phi_2 \phi_1^i = 0$$

and since  $x_1 \neq x_2$ ,

$$\phi_2 \phi_1^i = 0,$$

and  $\phi_1$  and  $\phi_2$  are orthogonal.

The following theorem is the principal result of this chapter and is of prime importance in the theory of quadratic forms.

**THEOREM 1.4:** If A is a real symmetric matrix, then there exists an orthogonal matrix P such that

$$P'AP = D$$

where D is the diagonal matrix of the characteristic roots of A.

**PROOF:** Let  $m(x)$  be the minimum function of A. Since the roots of the minimum equation of A are distinct

$$m(x) = (x_1 - x) \cdots (x_k - x)$$

where  $x_1, \dots, x_k$  are real and distinct.

Let  $S_1^i$  be the null space of  $x_1 - A$  and suppose the dimension of  $S_1^i$  is  $r_1$ . By Theorem 1.2,  $r_1$  linearly independent orthonormal vectors can be found which span  $S_1^i$ . Let these vectors be

$$\alpha_{i1}^i, \dots, \alpha_{ir_1}^i$$

all column vectors. Now, form the matrix P as follows:

$$P = (\alpha_{11}^1, \dots, \alpha_{1r_1}^1, \dots, \alpha_{k1}^k, \dots, \alpha_{kr_k}^k).$$

By Theorem 1.3, each vector of the space  $S_1^i$  is orthogonal to every vector of  $S_j^i$  where  $i \neq j$ , therefore, P is orthogonal.

Now

$$(x_1 I - A)\alpha_i^1 = 0$$

for every  $\alpha_i^1$  in  $S_1^1$  and

$$(x_1 I - A)\alpha_i^1 = x_1 I \alpha_i^1 - A \alpha_i^1 = 0,$$

and

$$A \alpha_i^1 = x_1 I \alpha_i^1.$$

Consequently,

$$A \alpha_{11}^1 = \alpha_{11}^1 x_1, \dots, A \alpha_{1r_1}^1 = \alpha_{1r_1}^1 x_1, \dots, A \alpha_{kr_k}^k = \alpha_{kr_k}^k x_k.$$

Combining these equations into one matrix equation,

$$AP = PD$$

where  $D$  is the diagonal matrix with the  $x_i$ 's on the diagonal.  
Since  $P$  is orthogonal,

$$P'AP = P'PD = D.$$

This is the desired transformation.

The next theorem is of value in proving one of the results of Chapter III.

**THEOREM 1.5:** If  $A$  is a real symmetric square matrix, then a necessary and sufficient condition that  $A$  be idempotent is that the characteristic roots,  $\lambda$ , of  $A$  be idempotent.

**PROOF:** 1. Suppose  $A$  is idempotent, i. e.,  $A^2 = A$ . By Theorem 1.4, there exists an orthogonal matrix  $P$  such that

$$P'AP = D$$

where  $D$  is a diagonal matrix of the characteristic roots of  $A$ .

Now, if

$$A^2 = A,$$

then

$$D^2 = D,$$

because

$$D = P'AP = P'A^2P = P'AAP = P'APP'AP = D^2.$$

Consequently, since  $D$  is diagonal,

$$\lambda^2 = \lambda.$$

2. Suppose the characteristic roots,  $\lambda$ , of  $A$  are idempotent.  
If  $D$  is the diagonal matrix of the characteristic roots, then

$$D^2 = D.$$

Then,

$$D = P'AP = D^2 = P'APP'AP = P'A^2P,$$

and from

$$P'AP = P'A^2P$$

it follows that

$$A = A^2.$$

The following results are important in the algebraic theory of analysis of variance. They enable the null space of an arbitrary symmetric matrix to be represented as the null space of a matrix whose rows are orthonormal.

THEOREM 1.6: If

$$\alpha_1, \dots, \alpha_s$$

are s linearly independent lxn vectors, then there exists orthonormal lxn vectors

$$\beta_1, \dots, \beta_s$$

such that for any j,  $1 \leq j \leq s$ , the null space of  $\alpha_j$  is equal to the null space of  $\beta_i$ ,  $i = 1, \dots, j$ .

PROOF: Define the  $\beta$ 's inductively as follows:

$$(1) \quad \beta_1 = \frac{\alpha_1}{\sqrt{\alpha_1 \alpha_1'}}$$

$$(2) \quad r_1 = \alpha_1 - \alpha_1(\beta_1' \beta_1 + \dots + \beta_{i-1}' \beta_{i-1}), \quad i = 1, \dots, s$$

$$(3) \quad \beta_1 = \frac{r_1}{\sqrt{r_1 r_1'}}$$

From (3) it follows immediately that each  $\beta_i$  is normal, i. e.,

$$\beta_i \beta_i' = 1.$$

Consider  $\beta_1 \beta_2'$ .

$$\beta_1 \beta_2' = \beta_1 \left[ \frac{\alpha_2 - \alpha_2(\beta_1' \beta_1)}{\sqrt{r_2 r_2'}} \right] = \frac{\beta_1 \alpha_2' - \beta_1 \beta_1' \beta_1 \alpha_2'}{\sqrt{r_2 r_2'}} = \frac{\beta_1 \alpha_2' - \beta_1 \alpha_2'}{\sqrt{r_2 r_2'}} = 0$$

By induction it follows that  $\beta_1 \beta_i' = 0$ ,  $i = 1, \dots, s$ . Similarly, it follows by induction that for any  $\beta_j$ ,  $\beta_j \beta_i' = 0$ ,  $j = 1, \dots, s$ ;  $i = 1, \dots, s$ ;  $i \neq j$ . Therefore, the  $\beta$ 's are orthonormal.

It follows immediately from the definition of the  $\beta$ 's that

if

$$\beta_1 \mu^* = \dots = \beta_j \mu^* = 0,$$

where  $\{\mu: \beta_i \mu^* = 0, i = 1, \dots, j\}$ , then

$$\alpha_1 \mu^* = \dots = \alpha_j \mu^* = 0.$$

Since

$$\alpha_i = \frac{\beta_i}{\frac{1 - (\beta_1^2 \beta_1 + \dots + \beta_{i-1}^2 \beta_{i-1})}{\sqrt{r_1 r_1}}}$$

it follows that if

$$\alpha_1 \mu^* = \dots = \alpha_j \mu^* = 0,$$

then

$$\beta_1 \mu^* = \dots = \beta_j \mu^* = 0.$$

**THEOREM 1.7:** If A and B are axn and bxn matrices such that the rank of  $\begin{bmatrix} A \\ B \end{bmatrix}$  is a+b, then there exists matrices C and D of dimension axn and bxn such that

$$(1) \quad CC' = I_a$$

$$DD' = I_b$$

$$CD' = O_{a,b}$$

(2) The null space of A is equal to the null space of C and the null space of  $\begin{bmatrix} A \\ B \end{bmatrix}$  is equal to the null space of  $\begin{bmatrix} C \\ D \end{bmatrix}$ .

**PROOF:** Applying Theorem 1.6 to the rows of  $\begin{bmatrix} A \\ B \end{bmatrix}$  gives a matrix  $\begin{bmatrix} C \\ D \end{bmatrix}$  such that

$$\begin{bmatrix} C \\ D \end{bmatrix} [C'D'] = I_{a+b},$$

yielding the relations (1).

By Theorem 1.6,

$$\{\mu: A\mu^* = 0\} = \{\mu: C\mu^* = 0\}$$

and

$$\{\mu: \begin{bmatrix} A \\ B \end{bmatrix} \mu^* = 0\} = \{\mu: \begin{bmatrix} C \\ D \end{bmatrix} \mu^* = 0\}.$$

## CHAPTER II

### SELECTED THEOREMS FROM THE THEORY OF MATHEMATICAL STATISTICS

Certain important results from the theory of mathematical statistics are presented in this chapter. The principal theorems are concerned with vectors whose elements are random variables having either the normal distribution or the  $\chi^2$  distribution. The result which shows that the distribution function is uniquely determined by the moment generating function is assumed throughout the chapter.

In order to present a clear picture of the situation some basic definitions are given.

DEFINITION: If  $x_1, \dots, x_n$  are random variables and

$$\xi = (x_1, \dots, x_n),$$

then the expected value of  $\xi$ , is defined to be

$$E(\xi) = [E(x_1), \dots, E(x_n)].$$

In general, the expected value of a matrix whose elements are random variables is defined to be the matrix of expected values:

$$E[(x_{ij})] = (E(x_{ij})).$$

DEFINITION: If  $x_1, \dots, x_n$  are random variables with means  $\mu_1, \dots, \mu_n$ , respectively, then the covariance between  $x_i$  and  $x_j$ ,  $i = 1, \dots, n$ ;  $j = 1, \dots, n$ , is

$$\sigma_{ij} = E[(x_i - \mu_i)(x_j - \mu_j)].$$



Let  $\xi = (x_1, \dots, x_n)$  and  $\mu = (\mu_1, \dots, \mu_n)$ . The symmetric matrix whose elements are the covariances  $\sigma_{ij}$  is called the covariance matrix of  $x_1, \dots, x_n$ , or  $\text{cov}(\xi)$ .

It is readily found that  $\text{cov}(\xi) = E(\xi' \xi - \mu' \mu)$ .

The immediate consequences of these definitions are given in the following theorems.

**THEOREM 2.1:** If  $x_1, \dots, x_n$  are random variables,

$$\xi = (x_1, \dots, x_n) \text{ and } \eta = (y_1, \dots, y_n)$$

and

$$\eta = \xi M$$

where M is a matrix of constants, then

$$E(\eta) = E(\xi)M.$$

**PROOF:** Let  $M = (a_{ij})$  where each  $a_{ij}$  is a constant.

By the hypothesis,

$$\eta = \xi M = \left( \sum_{i=1}^n x_i a_{i1}, \dots, \sum_{i=1}^n x_i a_{in} \right).$$

$$E(\eta) = E(\xi M) = E\left( \sum_{i=1}^n x_i a_{i1}, \dots, \sum_{i=1}^n x_i a_{in} \right) =$$

$$\left[ E\left( \sum_{i=1}^n x_i a_{i1} \right), \dots, E\left( \sum_{i=1}^n x_i a_{in} \right) \right] =$$

$$\left[ \sum_{i=1}^n E(x_i) a_{i1}, \dots, \sum_{i=1}^n E(x_i) a_{in} \right] =$$

$$E(\xi)M.$$

**THEOREM 2.2:** If  $x_1, \dots, x_n$  are random variables with means  $\mu_1, \dots, \mu_n$ , respectively, covariance matrix S and

$$\xi = (x_1, \dots, x_n), \mu = (\mu_1, \dots, \mu_n), \eta = (y_1, \dots, y_n) \text{ and}$$

$$\eta = \xi M$$

where M is a matrix of constants, then the  $\text{cov}(\eta)$  is

$$M' S M.$$

PROOF: The covariance matrix of  $\eta$  is given by

$$\begin{aligned} E(\eta'\eta - M'\mu'\mu M) &= E(M'\{\xi'\xi\}M - M'\mu'\mu M) = \\ E[M'(\{\xi'\xi\} - \mu'\mu)M] &= M'[E(\{\xi'\xi\} - \mu'\mu)]M = \\ &M'SM. \end{aligned}$$

The following paragraphs are devoted to a discussion of the multivariate normal distribution with special emphasis on the development of the moment generating function of this distribution. The definition of the density function of this important distribution is given first.

DEFINITION: Let  $x_1, \dots, x_n$  be random variables. If their probability density function has the form

$$f(x_1, \dots, x_n) = \frac{\sqrt{|A|}}{(2\pi)^{n/2}} \cdot e^{-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij}(x_i - \mu_i)(x_j - \mu_j)}$$

where the matrix  $A = (a_{ij})$  is positive definite, then  $x_1, \dots, x_n$  are said to have the multivariate normal distribution.

Three well-known results are given here without proof. They will simplify somewhat the development of the moment generating function of the multivariate normal distribution.

(1) If  $A$  is the matrix of the coefficients in the quadratic form of the multivariate normal density function, then  $A^{-1}$  is the covariance matrix of the random variables.

(2) If  $D$ , the diagonal matrix of the characteristic roots of a symmetric matrix  $A$ , is given by  $P'AP'$  (see Theorem 1.4), then  $D^{-1}$  is given by  $P'A^{-1}P$ .

$$(3) \quad \int_{-\infty}^{\infty} e^{-\frac{1}{2}cy^2 + ay} dy = \frac{\sqrt{2\pi}}{\sqrt{c}} \cdot e^{\frac{a^2}{2c}}$$

The importance of the moment generating function of the multivariate normal distribution warrants the detailed development given in the following theorem.

**THEOREM 2.3:** If  $\xi = (x_1, \dots, x_n)$  the components of which have the multivariate normal distribution with means  $\mu = (\mu_1, \dots, \mu_n)$ , then the moment generating function,  $m(t_1, \dots, t_n) = m(\tau)$ , of this distribution is given by

$$m(\tau) = e^{A\tau} e^{\frac{1}{2}\tau A^{-1}\tau}.$$

**PROOF:** To begin, suppose that  $E(\xi) = 0$ . Then,

$$m(\tau) = c \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{\tau \xi} e^{-\frac{1}{2} \xi A \xi} d\xi$$

where  $c = \frac{\sqrt{|A|}}{(2\pi)^{n/2}}$ . Now, let  $\eta = (y_1, \dots, y_n)$  and  $f = (z_1, \dots, z_n)$ .

By Theorem 1.4, there exists orthogonal matrix  $P$  such that

$$P^t A P = D$$

where  $D$  is the diagonal matrix of the characteristic roots,  $\lambda_i$ , of  $A$ .

Substituting  $\xi = \eta P$  and  $\tau = f P$  gives

$$\begin{aligned} & c \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{f P P^t \eta} e^{-\frac{1}{2} \eta P A P^t \eta} d\eta = \\ & c \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{f \eta} e^{-\frac{1}{2} \eta D \eta} d\eta = \\ & c \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{\sum_{i=1}^n z_i y_i} e^{-\frac{1}{2} \sum_{i=1}^n \lambda_i y_i^2} d\eta = \\ & c \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \lambda_1 y_1^2 + z_1 y_1} \dots e^{-\frac{1}{2} \lambda_n y_n^2 + z_n y_n} dy_1 \dots dy_n. \end{aligned}$$

Now, substituting

$$c = \frac{\sqrt{|A|}}{(2\pi)^{n/2}}$$

and applying remark (3) gives

$$\frac{\sqrt{|A|}}{(2\pi)^{n/2}} \cdot \frac{e^{-\frac{z_1^2}{2\lambda_1}} \sqrt{2\pi} \cdots e^{-\frac{z_n^2}{2\lambda_n}} \sqrt{2\pi}}{\sqrt{\lambda_1} \cdots \sqrt{\lambda_n}} =$$

$$e^{-\frac{1}{2} \sum_{i=1}^n \frac{z_i^2}{\lambda_i}} =$$

$$e^{-\frac{1}{2} \tau D^{-1} \tau}.$$

Then, applying remark (2),

$$e^{-\frac{1}{2} \tau P^* P A^{-1} P^* \tau} = e^{-\frac{1}{2} \tau A^{-1} \tau}.$$

Now, suppose  $E(\xi) = \mu$ , then  $m(\tau)$  is

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{t_1 x_1} \cdots e^{t_n x_n} e^{-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} (x_i - \mu_i)(x_j - \mu_j)} dx_1 \cdots dx_n =$$

$$e^{\mu_1 t_1} \cdots e^{\mu_n t_n} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} e^{t_1 (x_1 - \mu_1)} \cdots$$

$$e^{t_n (x_n - \mu_n)} e^{-\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n a_{ij} (x_i - \mu_i)(x_j - \mu_j)} dx_1 \cdots dx_n =$$

$$e^{\sum_{i=1}^n \mu_i t_i} e^{-\frac{1}{2} \tau A^{-1} \tau}.$$

Therefore,

$$m(\tau) = e^{\mu \tau} e^{-\frac{1}{2} \tau A^{-1} \tau}.$$

The following theorem is a basic result regarding the transformation of normally distributed random variables.

**THEOREM 2.4:** If  $x_1, \dots, x_n$  have the multivariate normal distribution,  $\xi = (x_1, \dots, x_n)$ ,  $\eta = (y_1, \dots, y_n)$ , and  $\eta = \xi M$

where M is a matrix of constants, then  $y_1, \dots, y_n$  have the multivariate normal distribution.

PROOF: Let  $\tau = (t_1, \dots, t_n)$ . Then

$$m(\tau) = E(e^{\tau \tau'}) = E(e^{f M \tau'}).$$

Let  $f' = M \tau'$ . Then

$$E(e^{f M \tau'}) = E(e^{f f'}).$$

Now, by Theorem 2.3,

$$E(e^{f f'}) = e^{\frac{1}{2} f S f'}$$

where S is the covariance matrix of  $x_1, \dots, x_n$ .

Substituting,

$$e^{\frac{1}{2} f S f'} = e^{\frac{1}{2} \tau M' S M \tau'}$$

and by Theorem 2.2,  $M' S M$  is the covariance matrix of  $y_1, \dots, y_n$ , so by Theorem 2.3 this is the moment generating function of the multivariate normal distribution and the proof is complete.

A relationship between the multivariate normal distribution and the  $\chi^2$  distribution is given in the following discussion. For the sake of completeness the  $\chi^2$  distribution is defined here.

DEFINITION: If the moment generating function of the random variable x is given by the quantity

$$(1 - 2t)^{-n/2}$$

then x is said to have the  $\chi_n^2$  distribution with n degrees of freedom.

The following well-known result, given without proof, is used in the proof of the next theorem: If  $x_1, \dots, x_n$  are independent and have the multivariate normal distribution with means zero and variances one, then the quantity

$$y = \sum_{i=1}^n x_i^2$$

has the  $\chi_n^2$  distribution.

**THEOREM 2.5:** If  $x_1, \dots, x_n$  have the multivariate normal distribution with means  $\mu_1, \dots, \mu_n$ , respectively, then the quadratic form

$$\sum_{i=1}^n \sum_{j=1}^n a_{ij} (x_i - \mu_i) (x_j - \mu_j)$$

appearing in the multivariate normal density function has the  $\chi_n^2$  distribution.

**PROOF:** Let

$$y = \sum_{i=1}^n \sum_{j=1}^n a_{ij} (x_i - \mu_i) (x_j - \mu_j)$$

and

$$\mathbf{f} = [(x_1 - \mu_1), \dots, (x_n - \mu_n)].$$

Then

$$y = \mathbf{f} \mathbf{A} \mathbf{f}'$$

where  $\mathbf{A}$  is the matrix of the quadratic form. Let  $\mathbf{f} = (z_1, \dots, z_n)$  and  $\mathbf{P}$  be an orthogonal matrix such that  $\mathbf{P} \mathbf{A} \mathbf{P}' = \mathbf{D}$ , the diagonal matrix of the characteristic roots,  $\lambda$ , of  $\mathbf{A}$ . (See Theorem 1.4.)

The moment generating function of  $y$  is given by

$$m(t) = E(e^{y t}) = E(e^{\mathbf{f} \mathbf{A} \mathbf{f}' t}).$$

Let  $\mathbf{f} = \mathbf{f} \mathbf{P}'$  and substitute, giving

$$\begin{aligned} E(e^{\mathbf{f} \mathbf{P}' \mathbf{A} \mathbf{P} \mathbf{f}' t}) &= E(e^{\mathbf{f} \mathbf{D} \mathbf{f}' t}) = \\ &E(e^{z_1^2 \lambda_1 t} \dots e^{z_n^2 \lambda_n t}) = \\ &E(e^{z_1^2 \lambda_1 t}) \dots E(e^{z_n^2 \lambda_n t}). \end{aligned}$$

Now

$$E(f) = E(f)P = (0, \dots, 0)$$

and

$$\text{cov}(f) = \left(\frac{1}{\lambda_1}, \dots, \frac{1}{\lambda_n}\right).$$

Hence, by the remark above,

$$\sum_{i=1}^n z_i^2 \lambda_i$$

has the  $\chi_n^2$  distribution.

But

$$\sum_{i=1}^n z_i^2 \lambda_i = f D f' = \{P D P' f'\} = \{A f'\} = y.$$

Consequently,  $y$  has the  $\chi_n^2$  distribution.

## CHAPTER III

### QUADRATIC FORMS IN NORMALLY DISTRIBUTED RANDOM VARIABLES

The distribution of quadratic forms involving normally distributed random variables is an important topic in the theory of mathematical statistics. Some of the results pertaining to the independence of quadratic forms and conditions under which these forms have the  $\chi^2$  distribution are given in this chapter.

Throughout this chapter  $\mathbf{f} = (x_1, \dots, x_n)$  will be a vector whose components,  $x_1, \dots, x_n$ , are independent normally distributed random variables with means zero and variances one. This fact will be stated in the succeeding theorems as follows:  $\mathbf{f}$  is a normal random vector with  $E(\mathbf{f}) = \mathbf{0}$  and  $\text{cov}(\mathbf{f}) = \mathbf{I}$ . The quadratic forms will be represented by  $Q = \mathbf{f} \mathbf{A} \mathbf{f}'$  where  $\mathbf{A}$  is an  $n \times n$  real symmetric matrix.

In order to prove the principal result of the chapter it is necessary to know the moment generating function of a quadratic form in normally distributed random variables. This development follows.

THEOREM 3.1: Let  $\mathbf{f}$  be a normal random vector with  $E(\mathbf{f}) = \mathbf{0}$  and  $\text{cov}(\mathbf{f}) = \mathbf{I}$ . Let  $Q = \mathbf{f} \mathbf{A} \mathbf{f}'$  be a quadratic form in the random variables. Then the moment generating function of  $Q$  is given by

$$|I - \lambda A|^{-1/2}.$$



PROOF: The density function for the random variables  $x_1, \dots, x_n$  is

$$\frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2} \mathbf{f} \mathbf{f}'}.$$

The moment generating function of  $Q$  is

$$\begin{aligned} E(e^{Q\mathbf{t}}) &= \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{\mathbf{f} \mathbf{A} \mathbf{f}' \mathbf{t}} e^{-\frac{1}{2} \mathbf{f} \mathbf{f}'} d\mathbf{f} = \\ &= \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \mathbf{f} \mathbf{f}'} + \mathbf{f} \mathbf{A} \mathbf{f}' \mathbf{t} d\mathbf{f} = \\ &= \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \mathbf{f} (\mathbf{I} - 2\mathbf{t} \mathbf{A}) \mathbf{f}'} d\mathbf{f}. \end{aligned}$$

Now, from the definition of the multivariate normal distribution,

$$\frac{|\mathbf{I} - 2\mathbf{t} \mathbf{A}|^{-\frac{1}{2}}}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \mathbf{f} (\mathbf{I} - 2\mathbf{t} \mathbf{A}) \mathbf{f}'} d\mathbf{f} = 1.$$

Consequently,

$$\begin{aligned} \frac{1}{(2\pi)^{n/2}} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} e^{-\frac{1}{2} \mathbf{f} (\mathbf{I} - 2\mathbf{t} \mathbf{A}) \mathbf{f}'} d\mathbf{f} = \\ |\mathbf{I} - 2\mathbf{t} \mathbf{A}|^{-\frac{1}{2}}. \end{aligned}$$

Let  $\lambda = 2\mathbf{t}$  and this becomes

$$|\mathbf{I} - \lambda \mathbf{A}|^{-\frac{1}{2}}.$$

The following theorem which gives conditions under which two quadratic forms are independent is the principal result of this chapter.

**THEOREM 3.2:** Let  $\mathbf{f}$  be a normal random vector with  $E(\mathbf{f}) = \mathbf{0}$  and  $\text{cov}(\mathbf{f}) = \mathbf{I}$ . Let  $Q_1 = \mathbf{f} \mathbf{A} \mathbf{f}'$  and  $Q_2 = \mathbf{f} \mathbf{B} \mathbf{f}'$  where  $\mathbf{A}$  and  $\mathbf{B}$  are symmetric  $n \times n$  matrices. Then a necessary and sufficient condition that  $Q_1$  and  $Q_2$  be independent is that

$$\mathbf{A} \mathbf{B} = \mathbf{0}.$$

PROOF: The moment generating functions are as follows:

$$(1) \quad E(e^{\frac{1}{2} \lambda Q_1}) = |I - \lambda A|^{-\frac{1}{2}}$$

$$(2) \quad E(e^{\frac{1}{2} \mu Q_2}) = |I - \mu B|^{-\frac{1}{2}}$$

$$(3) \quad E(e^{\frac{1}{2}(\lambda Q_1 + \mu Q_2)}) = |I - \lambda A - \mu B|^{-\frac{1}{2}}$$

Consequently, a necessary and sufficient condition for independence is that

$$|I - \lambda A| \cdot |I - \mu B| = |I - \lambda A - \mu B|$$

for all values of  $\lambda$  and  $\mu$ .

1. Suppose  $AB = 0$ . Now, the product of the determinants of two matrices is equal to the determinant of the product of the matrices, therefore,

$$|I - \lambda A| \cdot |I - \mu B| = |I - \lambda A - \mu B + \lambda \mu AB|.$$

Obviously,  $AB = 0$  implies the independence of  $Q_1$  and  $Q_2$  and the condition is sufficient.

2. Suppose  $Q_1$  and  $Q_2$  are independent. It must be shown that

$$AB = 0.$$

Now, a real orthogonal transformation can be found reducing  $Q_1$  to

$$d_1 x_1^2 + \dots + d_r x_r^2$$

where  $r$  is the rank of  $Q_1$ . Therefore, there exists an orthogonal matrix  $P$  such that

$$A = PLP' \text{ and } B = PMP'$$

where  $L$  and  $M$  are partitioned so as to separate the rows and columns into successive groups of  $r$  and  $n-r$  and are of the forms

$$L = \begin{bmatrix} D & O \\ O & O \end{bmatrix} \text{ and } M = \begin{bmatrix} E & C \\ C' & F \end{bmatrix}$$

where  $D$  stands for an  $r$ -rowed diagonal matrix having  $d_1, \dots, d_r$  in its diagonal,  $O$  for matrices whose elements are all zero, and  $E, C$

and  $F$  for arbitrary matrices of appropriate dimensions. Then,

$$|I - \lambda A| = |P(I - \lambda L)P^*| = |P| \cdot |I - \lambda L| \cdot |P^*| = |I - \lambda L|.$$

Similarly,

$$|I - \mu B| = |I - \mu M|$$

and

$$|I - \lambda A - \mu B| = |I - \lambda L - \mu M|.$$

Therefore,

$$|I - L| \cdot |I - M| = |I - L - M|.$$

Consequently, there exists quadratic forms  $Q_1^*$  and  $Q_2^*$  of variables which are normally and independently distributed with zero means and unit variances, and having matrices  $L$  and  $M$ , respectively, such that  $Q_1^*$  and  $Q_2^*$  are independent.

Now,

$$AB = PLMP^*$$

so the theorem will be proved if

$$LM = O.$$

Let

$$M = M_1 + M_2$$

where

$$M_1 = \begin{bmatrix} E & C \\ C^* & O \end{bmatrix} \text{ and } M_2 = \begin{bmatrix} O & O \\ O & F \end{bmatrix}.$$

Obviously

$$LM_2 = O.$$

It must be shown that

$$LM_1 = O.$$

Let

$$Q_2^* = Q' + Q''$$

where  $Q'$  and  $Q''$  are quadratic forms in the normal variables of  $Q_2^*$  and  $Q'$  and  $Q''$  have matrices  $M_1$  and  $M_2$  respectively. Now,  $Q_1^*$  and  $Q''$  are independent, so  $Q_1^*$  is independent of

$$Q_2^* - Q'' = Q'.$$

Therefore,

$$|I - \lambda L| \cdot |I - \mu M_1| = |I - \lambda L - \mu M_1|.$$

Note,

$$|I - \lambda L| = \prod_{i=1}^r (1 - \lambda d_i)$$

and

$$[I - \lambda L - \mu M_1] = \left| \begin{array}{ccc} I - \lambda D - \mu E & -\mu C \\ -\mu C' & I \end{array} \right|.$$

Substituting,

$$\prod_{i=1}^r (1 - \lambda d_i) \cdot |I - \mu M_1| = \left| \begin{array}{ccc} I - \lambda D - \mu E & -\mu C \\ -\mu C' & I \end{array} \right|.$$

Now, equating the coefficients of the terms containing the highest powers of  $\lambda$  on each side of the identity,

$$\prod_{i=1}^r d_i |I - \mu M_1| = |D|.$$

But,

$$\prod_{i=1}^r d_i |I - \mu M_1| = \prod_{i=1}^r d_i.$$

Therefore,

$$|I - \mu M_1| = 1.$$

Let  $\mu = \frac{1}{x}$  and substitute, giving

$$\left| I - \frac{1}{x} M_1 \right| = 1 \text{ or } |M_1 - xI| = (-x)^n.$$

Consequently, all the characteristic roots of  $M_1$  must be zero.

For a symmetric matrix the sum of the squares of the characteristic roots is equal to the sum of the squares of the elements since both are equal to the trace of the square of the matrix. Therefore,

$$M_1 = 0$$

and

$$LM_1 = 0$$

and the proof of the theorem is complete.

A result similar to Theorem 3.2 concerning a linear form and a quadratic form follows.

**THEOREM 3.3:** Let  $\mathbf{f}$  be a normal random vector with  $E(\mathbf{f}) = 0$  and  $\text{cov}(\mathbf{f}) = I$ . Let  $Q = \mathbf{f}'A\mathbf{f}$  and  $L = \mathbf{f}'\alpha$  where  $\alpha$  is a  $1 \times n$  vector. Then a necessary and sufficient condition that  $Q$  and  $L$  be independent is that

$$A\alpha' = 0.$$

**PROOF:** 1. Suppose that  $Q$  and  $L$  are independent. Then it follows that  $Q$  and  $L^2$  are independent.

Now,

$$L^2 = \mathbf{f}'\alpha'\alpha\mathbf{f}$$

is a quadratic form in the  $x_i$ 's and has matrix

$$B = \alpha'\alpha.$$

So, by Theorem 3.2,

$$AB = 0$$

and it follows immediately that

$$A\alpha' = 0.$$

2. Suppose that  $A\alpha' = 0$ . Then it follows that 0 is a characteristic root of  $A$  and that  $\alpha$  is a corresponding characteristic

vector. Let  $\lambda_2, \dots, \lambda_n$  be the remaining characteristic roots and  $\beta_2, \dots, \beta_n$  the corresponding characteristic vectors. Then

$$Q = \sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j = \sum_{i=2}^n \lambda_i (\beta_i' x)^2.$$

Since the  $\beta$ 's are orthogonal to  $\alpha$ , it follows that the linear combinations  $\beta_i' x$  are independent of  $\alpha' x$  and the proof is complete.

The remainder of this chapter is devoted to theorems concerning the distribution of quadratic forms.

**THEOREM 3.4:** Let  $\mathbf{f}$  be a normal random vector with  $E(\mathbf{f}) = 0$  and  $\text{cov}(\mathbf{f}) = \mathbf{I}$ . Let  $Q = \mathbf{f}'\mathbf{A}\mathbf{f}$ . Then a necessary and sufficient condition that  $Q$  have the  $\chi_r^2$  distribution is that the characteristic roots,  $\lambda_i$ , of  $\mathbf{A}$  be idempotent (0 or 1). The number of non zero roots is given by  $r$ , the rank of  $\mathbf{A}$ .

**PROOF:** 1. Suppose  $Q$  has the  $\chi_r^2$  distribution. By definition the moment generating function of  $Q$  is

$$(1 - 2t)^{-r/2}.$$

By Theorem 3.1 the moment generating function of any quadratic form is

$$|I - 2t\mathbf{A}|^{-1/2}$$

where  $\mathbf{A}$  is the matrix of the quadratic form. Therefore,

$$|I - 2t\mathbf{A}|^{-1/2} = (1 - 2t)^{-r/2}$$

and raising each side to the -2 power gives

$$|I - 2t\mathbf{A}| = (1 - 2t)^r.$$

Letting  $2t = \lambda$ , and substituting,

$$|I - \lambda\mathbf{A}| = (1 - \lambda)^r.$$

Consequently, it follows that the characteristic roots of  $A$  consist of  $r$  one's and  $(n - r)$  zero's.

2. Suppose the characteristic roots,  $\lambda$ , of  $A$  are idempotent. Since the rank of  $A$  is  $r$ , there are  $r$  non zero roots, and so

$$|I - \lambda A| = (1 - \lambda)^r.$$

Now, raising each side to the  $-1/2$  power and letting  $\lambda = 2t$ ,

$$|I - 2tA|^{-1/2} = (1 - 2t)^{-r/2}.$$

But, the left-hand side is the moment generating function of the quadratic form  $Q$  and the right-hand side is the moment generating function of the  $\chi^2_r$  distribution, consequently,  $Q$  has the  $\chi^2_r$  distribution.

**THEOREM 3.5:** Let  $Q = \{A\}$  where  $f$  is a normal random vector with  $E(f) = 0$  and  $\text{cov}(f) = I$ . Then  $Q$  has the  $\chi^2_r$  distribution if and only if the matrix  $A$  of the quadratic form is idempotent.

**PROOF:** 1. Suppose  $Q$  has the  $\chi^2_r$  distribution. Then, by Theorem 3.4, the characteristic roots,  $\lambda$ , of  $A$  are idempotent and

$$\lambda^2 = \lambda.$$

Consequently, by Theorem 1.5,  $A$  is idempotent.

2. Suppose  $A$  is idempotent. By Theorem 1.5 the characteristic roots of  $A$  are idempotent and it follows from Theorem 3.4 that  $Q$  has the  $\chi^2_r$  distribution.

**THEOREM 3.6:** If  $A_1, \dots, A_n$  are positive semi-definite symmetric matrices such that

$$A_1 + \dots + A_n = I$$

and if  $A_i$  and  $A_j$ ,  $i \neq j$ , are idempotent, then

$$A_i A_j = 0.$$

PROOF: Suppose  $A_i$  and  $A_j$ ,  $i \neq j$ , are idempotent. Choose  $P$  orthogonal such that  $PA_i P'$  is diagonal with the first  $r_i$  diagonal elements one, the rest zero. Since

$$PA_1 P' + \dots + PA_n P' = I$$

and each  $PA_i P'$ ,  $i = 1, \dots, n$ , is positive semi-definite, the first  $r_i$  diagonal elements of  $PA_j P'$  must be zero. Then since  $PA_j P'$  is idempotent, its first  $r_i$  rows, and hence columns, must all be zero, so that

$$0 = PA_i P' PA_j P' = A_i A_j.$$

THEOREM 3.7: Let  $A_1, \dots, A_n$  be positive semi-definite symmetric matrices such that

$$A_1 + \dots + A_n = I.$$

Let

$$Q_1 = \{A_1\}', \dots, Q_n = \{A_n\}'$$

where  $\{$  is a normal random vector with  $E(\{) = 0$  and  $\text{cov}(\{) = I$ . Then a necessary and sufficient condition that  $Q_i$  and  $Q_j$ ,  $i \neq j$ , be independent and have  $\chi^2$  distributions is that  $A_i$  and  $A_j$  be idempotent.

PROOF: 1. Suppose  $Q_i$  and  $Q_j$  are independent and have  $\chi^2$  distributions. Then, by Theorem 3.5,  $A_i$  and  $A_j$  are idempotent.

2. Suppose  $A_i$  and  $A_j$  are idempotent. By Theorem 1.8,

$$A_i A_j = 0,$$

and by Theorem 3.2,  $Q_i$  and  $Q_j$  are independent. It follows from Theorem 3.5 that  $Q_i$  and  $Q_j$  have  $\chi^2$  distributions.



## BIBLIOGRAPHY

- Aitken, A. C., Determinants and Matrices. New York: Interscience Publishing Company, 1939. 135 pp.
- Craig, Allen T., "A Note on the Independence of Certain Quadratic Forms," Annals of Mathematical Statistics, 14: 195-197, 1943.
- Cramér, Harald, Mathematical Methods of Statistics. Princeton: Princeton University Press, 1951. 575 pp.
- Hoel, Paul G., Introduction to Mathematical Statistics. New York: John Wiley and Sons, Inc., 1947. 258 pp.
- Hotelling, Harold, "Note on a Matrix Theorem of A. T. Craig," Annals of Mathematical Statistics, 15: 427-429, 1944.
- Kac, M., "A Remark on the Independence of Linear and Quadratic Forms Involving Independent Gaussian Variables," Annals of Mathematical Statistics, 16: 400-401, 1945.
- MacDuffee, Cyrus Colton, Vectors and Matrices. Menasha, Wisconsin: The Mathematical Association of America, 1943. 203 pp.
- Mood, Alexander McFarlane, Introduction to the Theory of Statistics. New York: McGraw-Hill Book Company, Inc., 1950. 433 pp.
- Ogawa, Junjiro, "On the Independence of Bilinear and Quadratic Forms," Annals of the Institute of Statistical Mathematics, 1: 49-57, 1949.
- Perlis, Sam, Theory of Matrices. Cambridge: Addison-Wesley Publishing Company, Inc., 1952. 237 pp.
- Wade, Thomas L., The Algebra of Vectors and Matrices. Cambridge: Addison-Wesley Publishing Company, Inc., 1951. 189 pp.